Lecture 7 Fast Johnson-Lindenstrauss

Let's fix a set $\mathcal{U} \subseteq \mathbb{R}^d$ of m points. JL tells us that a Gaussian matrix $X \in \mathbb{R}^{n \times d}$ satisfies

$$1-\varepsilon \leq \frac{\|Xu - Xv\|_2}{\|u-v\|_2} \leq 1+\varepsilon \quad \forall u,v \in \mathcal{U}$$

with probability $1-\delta$, as long as

$$\boxed{n = \Omega\left(\frac{\log\left(\frac{m}{\delta}\right)}{\varepsilon^2}\right)}$$

The cost of forming Xu is

$O(dn)$ flops.

We will prove that there exists a matrix $L \in \mathbb{R}^{n \times d}$ that satisfies the conclusions of JL but forming Lu costs

$O(d\log(d) + n\log(dm))$ flops.

__First attempt__: Let's try to construct a super sparse matrix with a single nonzero entry per row.

For each i, let

$$L_i = \sqrt{\frac{d}{n}}\, e_j \quad \text{w.p. } \frac{1}{d}$$

__E.g.__:

$$\begin{array}{cc} & d \\ n & \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}} \end{array}$$

Let's check the norm in expectation:

$$\mathbb{E}\|Lu\|_2^2 = \sum_i \mathbb{E}(L_i u)^2$$

$$= \sum_i \sum_j \frac{d}{n} \cdot u_j^2 \cdot \frac{1}{d}$$

$$= \frac{1}{n} \sum_i \sum_j u_j^2 = \|u\|_2^2$$

So $\|Lu\|_2^2$ has the right mean, but does not concentrate well for example when $u = e_1$. Indeed $Le_1 = 0$ with probability

$$p = \left(\frac{d-1}{d}\right)^n = \left(1 - \frac{1}{d}\right)^n$$

So $\log(p) = n \log\left(1 - \frac{1}{d}\right) = \frac{n}{d} + o\left(\frac{1}{d}\right)$

So need $n \gtrsim d$ to have $p < \frac{1}{2}$ for example.

On the other hand if u is not aligned with the axes Lu does concentrate well.

<u>Lemma</u>: Let $u \in \mathbb{R}^d$ be such that

$$\|u\|_\infty^2 \leq \lambda \cdot \|u\|_2^2$$

Then $P\left[\left|\|Lu\|_2^2 - \|u\|_2^2\right| \geq \varepsilon \|u\|_2^2\right] \leq 2\exp\left(-\frac{\varepsilon^2 n}{d^2 \lambda^2}\right)$

<u>pf</u>: Observe $\|Lu\|_2^2 = \sum_i (L_i u)^2$

and $(L_i u)^2 = \frac{d}{n} u_j^2$ w.p. $\frac{1}{d}$

Therefore $(L_i u)^2$ is $\frac{d\|u\|_\infty^2}{2n}$ - subGaussian

Hoeffding implies $\|Lu\|_2^2$ is $\frac{d\|u\|_\infty^2}{2\sqrt{n}}$ subGaussian and therefore

$$P\left[\left|\|Lu\|_2^2 - \|u\|_2^2\right| \geq \varepsilon \|u\|_2^2\right]$$

$$\leq 2\exp\left(-\frac{\varepsilon^2 \|u\|_2^4 n}{d^2 \|u\|_\infty^4}\right) \leq 2\exp\left(-\frac{\varepsilon^2 n}{d^2 \lambda^2}\right)$$

$\blacksquare$

Idea: Apply first a random matrix R to u_i so that Ru_i is not alligned with coordinate axes. Specifiall we would like

Incoherence

$$\|Ru_i\|_\infty \approx \frac{1}{\sqrt{d}} \|Ru_i\|_2$$

in order to apply the lemma.

But R should be such that Ru is easy to compute.

Defn: A Hadamard matrix $H \in \mathbb{R}^{d \times d}$ is one that is orthogonal and has all entries in $\left\{-\frac{1}{\sqrt{d}}, \frac{1}{\sqrt{d}}\right\}$

Ex: For $d=2$, have

$$H_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

45° degree rotation

For powers $d = 2^k$ we may set

$$H_d = \begin{bmatrix} H_{d/2} & H_{d/2} \\ H_{d/2} & -H_{d/2} \end{bmatrix} \Big/ \sqrt{2}$$

<u>Lemma</u>: H_d is Hadamard.

<u>pf</u>: By induction $H_d(i,j) = \pm \frac{\sqrt{2}}{\sqrt{d}} \Big/ \sqrt{2} = \pm \frac{1}{\sqrt{d}}$

and direct computation shows $H_d^T H_d = I$ $\blacksquare$

Note to compute $H_d u$, we may write

$$H_d \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} H_{d/2}(u_1 + u_2) \\ H_{d/2}(u_1 - u_2) \end{bmatrix}$$

So can proceed inductively. In each step we need to compute $u_1 + u_2$, $u_1 - u_2$ and there are $\log(d)$ steps. So total cost is $d \log(d)$.

We'll introduce randomness by letting D be a diagonal matrix with random sign $\varepsilon_i = \{-1, +1\}$ entries. Set $R = HD$. Can check that R is also Hadamard.

<u>Lemma</u>: For any u set $v = HDu$. Then

$$P\left[\frac{\|v\|_\infty}{\|v\|_2} \geq \sqrt{\underbrace{\frac{2\log\left(\frac{4d}{\delta}\right)}{d}}_{\lambda}}\right] \leq \frac{\delta}{2}$$

pf: Observe $\|v\|_2 = \|u\|_2$ and

$$v_i = \sum_j \varepsilon_j H_{ij} u_j$$

So by Hoeffding v_i is subGaussian with param $\sqrt{\sum_j H_{ij}^2 u_j^2} = \frac{\|u\|_2}{\sqrt{d}}$.

Therefore

$$P\left[|v_i| \geq \sqrt{\lambda}\,\|u\|_2\right] \leq 2\exp\left(-\frac{\lambda d}{2}\right) = \frac{\delta}{2d}$$

Taking union bound over i completes the pf $\blacksquare$

Finally putting the two lemmas together gives the thm.

Thm (Fast JL Transform)

Define the matrix

$$X = \underbrace{L}_{\text{supersparse}} \cdot \underbrace{H}_{\text{Hadamard}} \cdot \underbrace{D}_{\text{random diagonal}}$$

Then with probability $1-\delta$, we have

$$1-\varepsilon \leq \frac{\|Xu - Xv\|_2}{\|u-v\|_2} \leq 1+\varepsilon \quad \forall u,v \in \mathcal{U}$$

provided

$$n \geq \frac{2\log\left(\frac{4m^2}{\delta}\right)\log\left(\frac{4dm^2}{\delta}\right)}{\varepsilon^2}.$$

Forming Xu requires $d\log(d)+n$ flops.

Note this n is worse than n from JL by a factor of $\log(dm)$.

Pf: For any u, with prob $1-\frac{\delta}{2m^2}$ w.r.t. D we have

① $$\frac{\|HDu\|_\infty^2}{\|HDu\|_2^2} \leq \lambda = \frac{2\log\left(\frac{4dm^2}{\delta}\right)}{d}$$

Conditioned on this event E, with probability $p = 1-2\exp\left(-\frac{\varepsilon^2 n}{d^2\lambda^2}\right)$ we have

② $$\Big|\underbrace{\|LHDu\|^2}_{\|Xu\|^2} - \underbrace{\|HDu\|_2^2}_{= \|u\|_2^2}\Big| \leq \varepsilon \underbrace{\|HDu\|_2^2}_{= \|u\|_2^2}$$

Note $p = 1-2\exp\left(-\frac{\varepsilon^2 n}{2\log\left(\frac{4dm^2}{\delta}\right)}\right) = 1-\frac{\delta}{2m^2}$

Let E be the event where ① holds and F the event that ② holds. Then

$$\begin{aligned} P[E \wedge F] &= P[F \mid E] \cdot P[E] \\ &\geq \left(1-\frac{\delta}{2m^2}\right)\cdot\left(1-\frac{\delta}{2m^2}\right) \\ &\geq 1-\delta/m^2 \end{aligned}$$

Take union bound over $u = x-y$ for $x, y \in \mathcal{U}$ ☒